NASA CONTRACTOR REPORT

NASA CR-2471

NASA CR-2471

RELIABILITY OF LABORATORY TESTS OF VSTOL AND OTHER LONG-DURATION NOISES

by K. D. Kryter, D. J. Peeler, M. E. Dobbs, and J. S. Lukas

Prepared by
STANFORD RESEARCH INSTITUTE
Menlo Park, Calif. 94025
for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • NOVEMBER 1974

1. Report No. NASA CR-2471	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle RELIABILITY OF LABORATORY TESTS OF VSTOL AND OTHER LONG-DURATION NOISES		5. Report Date November 1974
		6. Performing Organization Code
7. Author(s) K. D. Kryter, D. J. Peeler, M. E. Dobbs and J. S. Lukas		8. Performing Organization Report No.
9. Performing Organization Name and Address Stanford Research Institute 333 Ravenswood Avenue Menlo Park, California 94025		10. Work Unit No.
		11. Contract or Grant No. NAS1-11602
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		13. Type of Report and Period Covered Contractor Report
		14. Sponsoring Agency Code

15. Supplementary Notes

Final Report

16. Abstract

Paired-comparison and magnitude estimations of the subjective noisiness or unacceptability of noise from fixed wing jet aircraft and simulated noise of VSTOL aircraft were obtained from groups of subjects given somewhat different instructions These results suggest that VSTOL noises can be evaluated in terms of their noisiness or unwantedness to people with reasonable accuracy by units of the physical measures designated as PNdBM, with or without tone corrections, and dBD_2. Also, that consideration should be given to the use of D_2 as an overall frequency weighting function for sound level meters instead of the presently available A weighting. Two new units of noise measurement, PLdB and dB(E), recently proposed by S. S. Stevens for predicting subjective noisiness were found to be less accurate than PNdBM or dBD_2 in this regard.

17. Key Words (Suggested by Author(s)) VSTOL NOISE; SLM Frequency Weighting; Psychoacoustic Tests; Aircraft Noise Evaluation		18. Distribution Statement Unclassified - Unlimited STAR Category: 04	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 43	22. Price* $3.25

*For sale by the National Technical Information Service, Springfield, Virginia 22151

CONTENTS

RELIABILITY OF LABORATORY TESTS OF VSTOL AND OTHER LONG-DURATION NOISES

K. D. Kryter, D. J. Peeler, M. E. Dobbs and J. S. Lukas
Stanford Research Institute, Menlo Park, California 94025

INTRODUCTION

The specification of maximum limits allowable for noise from aircraft and the noise limits to be allowed in communities are based on methods of noise measurement that take into account the spectral and temporal characteristics of the noise--the so-called Effective Perceived Noise Level (EPNL) either in EPNdB, EdBA, $EdBD_2$, etc. units. These units of noise measurement have been developed largely through subjective judgment tests conducted in laboratories.

One of the perplexing problems with these noise measurement evaluation procedures is the inconsistency in the results of some laboratory tests when the noise stimuli to be judged varied significantly with respect to both spectral content and duration or temporal pattern. The difficulty, however, may be more related to the experimental procedures followed in the laboratory than in how predictive the methods would be for reactions to the noises in "real-life". For example, in recent tests (Ref. 1) conducted to evaluate the effectiveness of the nacelle-noise reduction (Ref. 2 and 3) it was found that the units of noise measurement that take the duration of the noises into account do no better (even slightly worse because of the greater unreliability introduced by the increased number of physical measures required for obtaining effective values) in predicting the subjective judgments than do those units of measurement that reflected only the maximum level reached during the noise occurrence. This would be contrary to other test results (Ref. 4) and the common observation that the longer the duration of a noise the more objectionable it is. Does this mean, that noise measurements, at least for some classes of noises, should not utilize the duration information? The answer is probably no, because the noises in these tests were all of about

equal duration and, accordingly, the maximum intensity level and spectrum of each noise solely determined its relative judged noisiness.

Inasmuch as decisions regarding the use of particular methods of noise measurement and the modification of these methods and their standardization on an industry-government-wide basis depend to a large extent on laboratory test results, it is important to continue to verify and upgrade the methods used in these tests. Also, the noise of VSTOL aircraft represents a relatively new type of aircraft noise that differs from the noise of present-day fixed-wing aircraft with regard to both spectral content and duration. Accordingly, tests were conducted to study: (1) the judged annoyance effect of VSTOL noise in comparison with other present-day noises, (2) the reliability of the research methods used in these judgment tests; and (3) the relative accuracy of various older and two newer units recently proposed by S. S. Stevens (Ref. 5) of noise measurement in predicting the subjective perceived noisiness or unacceptableness of aircraft or other complex noises.

PROCEDURE

Acoustic Environment. All tests were conducted in an anechoic chamber which had 21-inch long fiberglass wedges on all six surfaces (see Fig. 1). Measured from the tips of the wedges the internal dimensions of the anechoic chamber were 8.5 by 17.75 by 8 feet. The noises to be judged were presented via two Altec-Lansing A7-500 speaker systems each driven by an 80 watt McIntosh power amplifier. Conventional playback circuitry was employed with the exception of artificial quieting of the system noise between stimulus presentations and the use of an equalization network designed to provide as flat as possible frequency response at the listener positions within the room. A block diagram, with manufacturer's name and model number of commercial equipment used specified, is provided in Fig. 2.

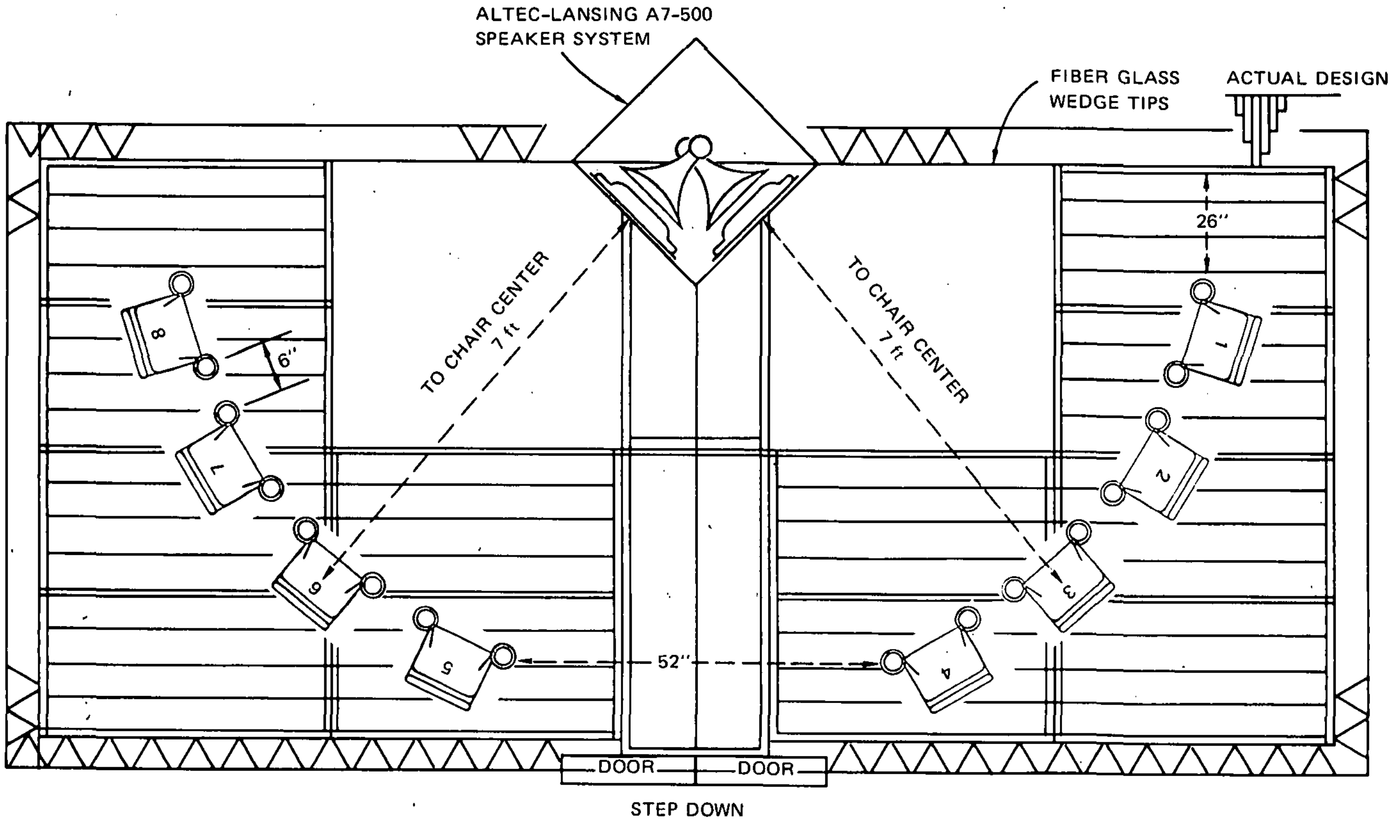
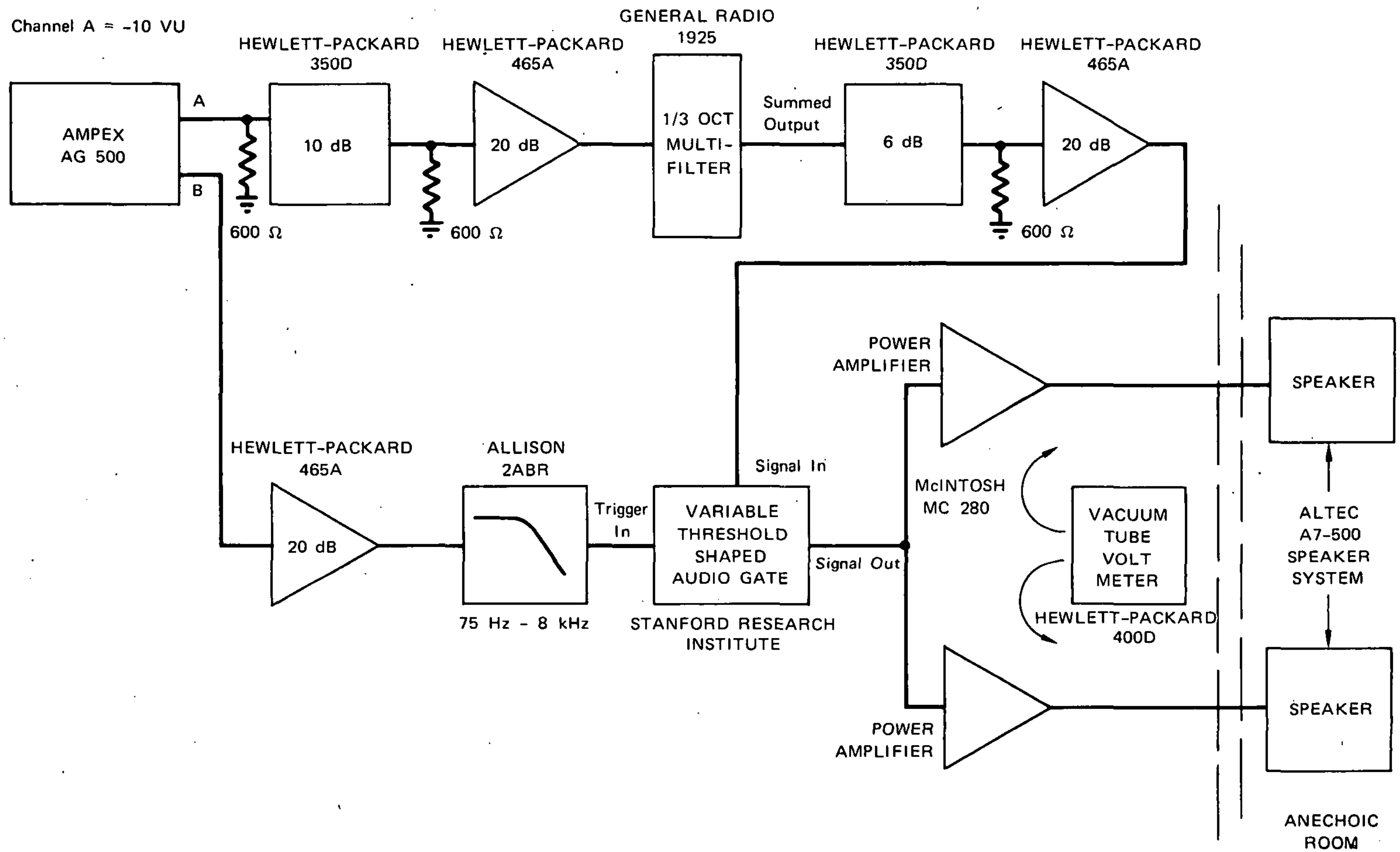


FIGURE 1. Showing anechoic chamber and location of subjects chairs and loudspeakers.



GR multifilter attenuators set to equalize anechoic room characteristics.
Amplifier and attenuator settings shown will produce 90 dBA at calibrate-microphone position in room using calibration noise on stimulus tape.

FIGURE 2 BLOCK DIAGRAM OF PLAY-BACK EQUIPMENT USED FOR TESTS

Each speaker system was directed at four subjects seated in an arc of radius of 8-1/2 feet. The chord of each arc was approximately 5 feet. The sound pressure level of octave bands of noise with center frequencies ranging from 63 to 8000 cycles varied by less than $\pm$ 2-1/2 dB at any listener position. A low-pass filter with 3 dB downpoint at 8000 Hz was used to minimize tape hiss.

Physical Analysis. Physical measures of noises were computed from one-third octave band sound pressure levels sampled and averaged over 1/2 second time intervals. A General Radio Type 1921 Real-Time Analyzer was used to produce, each 1/2 second, sound pressure level measurements in 24 one-third octave bands covering the frequency range 50 to 10,000 Hz. These data were recorded and processed in digital form. The end results of the analysis include the time-histories of sound pressure levels in each of the 24 bands and the so-called maximum (Max PNL) and effective levels (EPNL) of various weighted measures dBA, dBC, dBD_2, dBE, PLdB, PNdB, PNdBM, PNdB and PNdBM corrected for tonal content by two procedures and designated by the subscripts t1 and t2. These units and related frequency weightings and calculation procedures are given in detail in Refs. 3, 4, and 5 and are summarized in Table 1 and Figure 3.

Noise Stimuli. The various noise stimuli used in the judgment tests are described in Table 2. It is seen in Table 2 that the so-called VSTOL noises were actually simulated to have the spectra and approximate durations believed to be typical for such aircraft. All the noises were recorded onto a master tape with the same dBD_2 peak level. The relative intensity levels of the test items were appropriately varied by means of an attenuator during re-recording onto test tapes from the master tape. These test tapes were then played via loudspeakers to the listeners in the anechoic chamber. The equipment used for the making of the test tapes is shown in Figure 4.

Table 1 UNITS OF PHYSICAL NOISE MEASUREMENT

I	dBA, dBD_2, dBC, and dBE are sound level meter, with specified weightings (see Fig. 3) and meter action set on "slow". A and C weighting (dBA,dBC) and "slow" meter action are defined in Ref. 6. D_2 weighting (dBD_2) and PNdB and PNdBM with and without pure-tone corrections (t_1 and t_2) are defined in Ref 4. PLdB and E weighting (dBE) are defined in Ref. 5.
II	PNL is the level for each of the above units present in each successive half-second interval of a noise occurrence. Max PNL is the maximum level reached on a sound level meter overall weighted frequencies or the maximum PNdB, PNdBM (t_1) (t_2) and PLdB level reached during successive half-second intervals of a noise occurrence.
III	EPNL is taken as the integration on a 10 $\log_{10}$ basis of the half-second PNL values present between the 10 dB downpoints from the half-second interval in which the max PNL occurred.

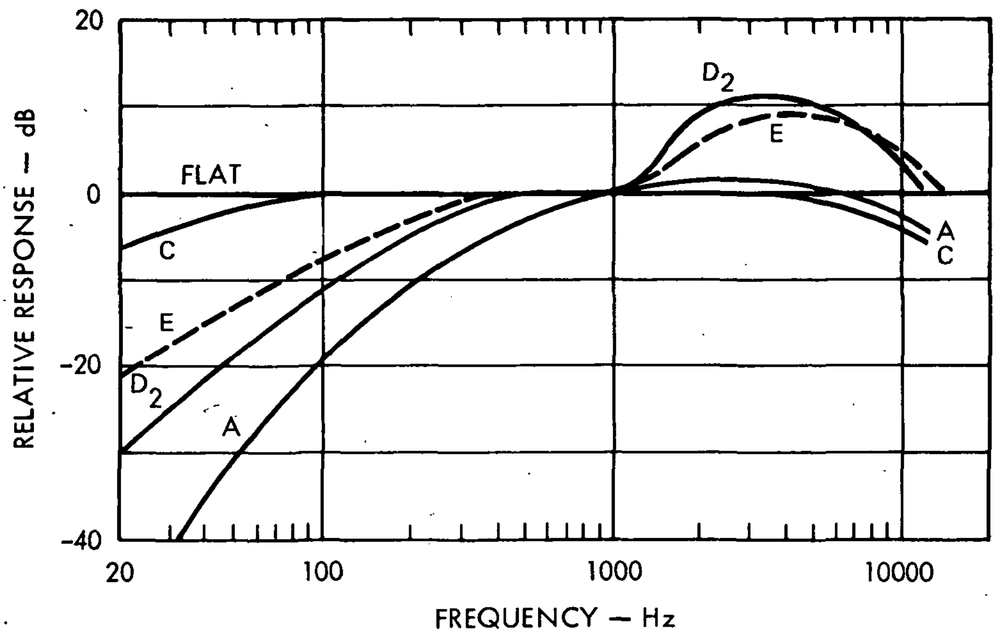


FIGURE 3. Showing frequency weightings applied to overall sound level measurements (see ref. 4, 5, and 6).

Table 2 NOISES USED IN PAIRED-COMPARISON AND MAGNITUDE ESTIMATION TESTS

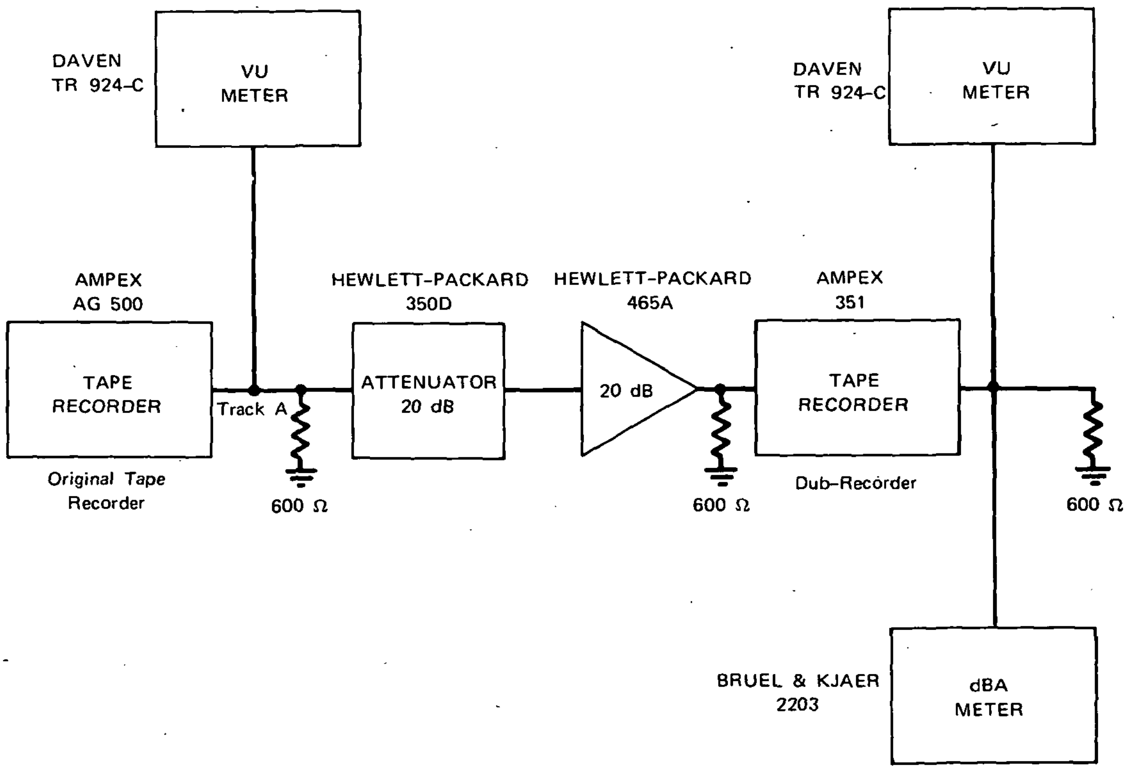
	Noise	Nominal Max Level	Duration to 10 dB Downpoints
Standard 1 (S1)	Pink Noise	80 dBA	4 secs.
Standard 2 (S2)	Pink Noise	73 dBA	9 secs.
		Presented in 5 Steps, 4 dB Apart	
"VSTOL" 1 (V1)*	737 Takeoff at 1/2 Speed Recording	72-88 dBA	8.7 secs.
"VSTOL" 2 (V2)*	737 Landing at 1/2 Speed Recording	70-86 dBA	6.0 secs.
"VSTOL" 3 (V3)*	737 Flyby at 1/2 Speed Recording	72-88 dBA	16.8 secs.
"VSTOL" 4 (V4)*	Vertol Simulation Recording	70-86 dBA	15.4 secs.
"VSTOL" 5 (V5)*	Vertol Simulation Recording	70-86 dBA	18.8 secs.
F1	737 Takeoff	72-88 dBA	5.2 secs.
F2**	737 Landing	70-86 dBA	3.4 secs.
F3	737 Flyby	72-88 dBA	16.9 secs.
F4	747 Takeoff	72-88 dBA	6.2 secs.
F5	747 Landing	70-86 dBA	7.2 secs.
F6	DC-8 Landing (Nacelle Treated)	72-88 dBA	9.6 secs.
F7	DC-8 Landing (Nacelle Untreated)	66-82 dBA	6.7 secs.
AC	Recording of Room Air-Conditioner	70-86 dBA	9.3 secs.

Note 1. S1 and S2 shaped at rate of 20 dB/.5 sec. below max level; at max level for 1.5 sec.

Note 2. All noises "faded" at rate of 30 dB/.5 sec. when level reaches 15 dB below max dBA level.

*Recordings furnished by NASA Langley Research Center.

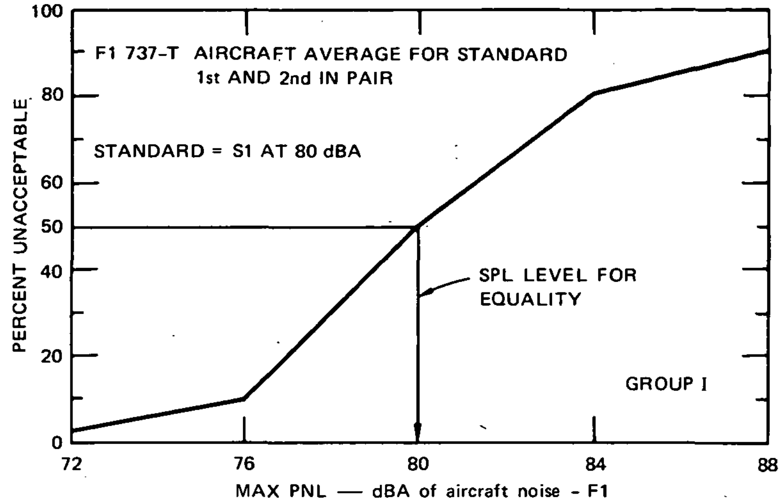
**The F2 noise had atypical "skirts" that caused difficulties in making certain physical measurements and was omitted for this reason from further data analysis.



Both tape recorders aligned with Ampex alignment tape number 01-31321-01 for equal VU output (play and record). After alignment, exact level dub is obtained with attenuator and amplifier settings as shown.

FIGURE 4. BLOCK DIAGRAM FOR DUBBING STIMULUS PRESENTATION TAPES

Paired Comparison Test. Each of the aircraft noises and the room air-conditioner noise were paired with each of the two standards to form a pair of noises. The subjects were asked (see Appendix A) to judge which in each pair they considered the more unacceptable, bothersome or annoying. The levels of the standard pink-noise (the output of an electronic random noise generator shaped to have a low frequency roll-off below 63 Hz of 3 dB per octave and a high frequency roll-off above 500 Hz of 6 dB per octave) and the other comparison noises were presented at the levels indicated on Table 2. Each comparison noise was paired at each indicated level twice with each of the standards, once occurring first in the pair and once second in the pair. The percentage of subjects who judged each of the comparison noises at each of its levels and in each order of presentation (when preceding and when following the standard noise in a pair) was placed on a graph showing percentage of subjects plotted against the sound pressure level of the comparison noise. The level, as determined from the resulting curve, at which 50 percent of the subjects would indicate the comparison noise was the more unacceptable (or 50 percent would indicate the standard to be the more unacceptable) was taken as the level of the noise as measured by a given unit of physical sound measurement that provided subjective equality with the standard. The values found for the two orders of presentation were average to provide an answer presumably free of the "time" error often present in such judgments due to the order in which a noise appears in each pair. An example of a paired-comparison test function for aircraft noise F-1 is given in the upper graph in Figure 5. Fifty pairs of noises, requiring about 20 minutes, were presented to the subjects in a single session with a minimum of 10 minutes rest between sessions. The pairs of noises were recorded with a 4 second pause between pairs and about 1 second pause between noises within a pair. Every 5th pair was preceded with a pair number announcement (recorded on the test tape) and a weak intensity beep tone separated the other pairs. The sequencing of comparison noises and levels was randomized on the test tapes.



EXAMPLE OF PAIRED-COMPARISON FUNCTION

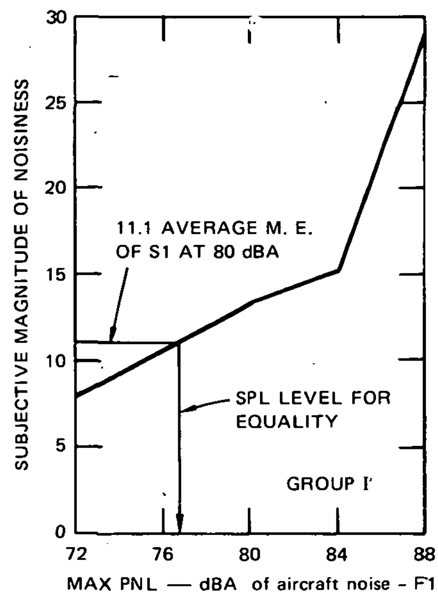


FIGURE 5. Examples of graphs developed from Paired-Comparison (upper) and Magnitude-Estimation (lower) test results.

Magnitude Estimation Test. Each of the noises at each of the levels used in the paired-comparison test was re-recorded in a random order onto a master test tape with about 3 seconds between noises, with 81 noises on a given tape. However, the first noise on each test tape was the standard pink noise S1 presented at a maximum level of 80 dBA. The subjects were instructed to ascribe to the magnitude of relative unacceptableness of that noise the number 10, and to judge each succeeding noise in relation to the standard and assign an appropriate number to it, e.g., if the second sound appeared to be twice as noisy or unacceptable it was to be given the number 20. The specific instructions appear in Appendix A.

The average for all subjects of a particular group of these magnitude judgments for a given noise were then plotted against the level, as measured by a given unit of physical noise measurement, at which the noise in question was presented. The standard noise was presented at its specified level 5 times during the course of 1 test tape of 81 noises and the average numerical magnitude of these five ratings was noted.

Using the average numerical magnitude given the standard, the graphs prepared for each of the other noises was then entered to determine the physical noise levels required, using a particular unit of physical noise measurement of each of the noises required, to achieve the equal numerical magnitude of subjective noiseiness. An example of how the magnitude estimation data were interpreted is given in the lower graph of Figure 5.

Instructions to the Subjects. Two different sets of instructions were prepared for both the paired-comparison test and the magnitude estimation test, as shown in Appendix A. One set was relatively detailed and repetitious with the intent of making the subjects concentrate and consider the whole noise occurrence and not just the peak levels that occurred in each noise. The appropriate instructions were repeated at the beginning of each rest period between sessions.

The second set of instructions were abbreviated as much as thought possible and given but once at the start of each type of test procedure.

Subjects. Three groups of 24 subjects each were selected. The subjects consisted of 11 male and 25 female college students and 9 housewives, all of whom reported they had no hearing difficulties. Group I was given the paired-comparison test first and about 2 weeks later, the magnitude-estimation test. The "long" form of instructions was used with Group I as with Group II. However, Group II received the magnitude-estimation test first and the paired-comparison test second about 1 week later. Group III received the short form of the instructions, the paired-comparison test first and the magnitude-estimation test 2 days later.

RESULTS AND DISCUSSION

Magnitude Estimation. The standard deviation(S.D.)statistic is used as a means of evaluating the accuracy with which each of the units of physical noise measurement predicted the magnitude estimation judgments made by the three groups of subjects. Normal probability statistics have been commonly used in the analysis of this type of judgment data (Refs. 1, 2, 3, 4, 5, 6, 7, 8, see particularly Ref. 2). The results are shown in Table 3.

These standard deviations are calculated by taking the square root of the average of the sum of the squared differences between: (1) the average level, as measured by a given physical unit required of each of the noises in order that they each be judged to have the same subjective magnitude; and (2) the level of each individual noise when judged to have the same magnitude (see Fig. 5). In formula, this is as follows:

$$\text{S.D.} = \sqrt{\frac{\Sigma\ (X-M)^2}{N-1}}$$

If there were perfect agreement between the physical unit of measurement and the subjective judgments, the physical levels would, of course, have

Table 3 MAGNITUDE ESTIMATION

Standard deviations of differences between physical levels of the noises and the mean of the distribution of the levels that elicit an equal subjective magnitude of acceptability.

Max PNL	Group			
	I	II	III	Average
dBC	5.51	5.42	5.41	5.45
dBA	2.21	2.02	2.20	2.14
dBD_2	2.20	1.53	2.03	1.92
dBE	2.71	2.11	2.39	2.40
PLdB	2.45	1.73	2.15	2.11
PNdB	2.21	1.48	1.87	1.85
PNdBM	2.22	1.40	2.08	1.90
$PNdBM_{t1}$	2.81	2.26	2.49	2.52
$PNdBM_{t2}$	2.80	1.50	1.63	1.98
$PNdB_{t2}$	1.98	1.58	1.74	1.77

EPNL	Group			
	I	II	III	Average
EdBC	5.73	6.00	5.60	5.78
EdBA	2.65	2.91	3.01	2.86
$EdBD_2$	1.89	2.00	2.26	2.05
EdBE	3.26	3.13	3.08	3.16
EPLdB	2.68	2.69	2.68	2.68
EPNdB	2.47	2.48	2.39	2.45
EPNdBM	2.30	2.12	2.39	2.27
$EPNdBM_{t1}$	1.80	1.72	1.86	1.79
$EPNdBM_{t2}$	1.80	1.75	1.98	1.84
$EPNdB_{t2}$	1.85	2.00	1.75	1.87

identical values and the distribution of differences from the average would be zero. This statistic can be presumed to show the relative accuracy, in dB, with which the different physical measures will predict subjective judgments of different noises. For example, it is seen in Table 3 that 67% of the noise having Max dBC levels that differed by as much as 10.90 dB (a plus or minus standard deviation of 5.45 dB would encompass in normal distributions about 67% of all cases or the noises) could be judged as subjectively equal; this percentage of the noises would be judged as equal when their Max D_2 values were within a range of but 3.84 dB (plus or minus a standard deviation of 1.92 dB).

It is clear from Table 3 that the results for the three groups are reasonably consistent with each other. Also, that the Max PNL units of physical measurement are at least as good as the EPNL units. This finding, and a comparison of the relative accuracies of the different physical units of measurement in predicting the subjective judgments of these noises will be discussed later, after presentation of the results of the paired-comparison tests.

Paired-Comparison Test. The standard deviations are shown in Table 4 of the distribution of differences for each of the groups, between the level of each of the standards, S1 and S2, and each of the comparison noises when judged to be equally unacceptable and when measured in terms of the various physical units. Again, as with the magnitude estimation tests, the Max PNL units exhibit standard deviations or errors of prediction that are at least as small, on the average, as those for EPNL. Also, it is clear from Table 4 that there are no large apparent differences among the results found when the longer duration (9 secs.) standard reference noise S2 was used compared to those obtained when the comparisons were made against the shorter (4 secs.) standard noise (S1).

Table 4 - Part 1 PAIRED-COMPARISON TEST - MAX PNL

Standard deviation of differences between the physical level of a standard noise and of comparison noises that are found to be subjectively equal in acceptability according to Paired-Comparison tests.

Standard 1				
	Group			
Max PNL	I	II	III	Average
dBC	4.91	6.56	5.40	5.62
dBA	1.94	3.24	2.07	2.42
dBD_2	1.63	2.18	1.15	1.65
dBE	2.29	3.02	1.92	2.41
PLdB	1.85	2.72	1.55	2.04
PNdB	1.64	2.29	1.31	1.75
PNdBM	1.62	2.39	1.23	1.75
$PNdBM_{t1}$	2.17	1.82	2.07	2.02
$PNdBM_{t2}$	2.17	2.20	1.07	1.81
$PNdB_{t2}$	2.37	2.91	1.66	2.31

Standard 2					
	Group				
Max PNL	I	II	III	Average	Aver. of S1 and S2
dBC	5.29	7.32	7.31	6.64	6.13
dBA	2.23	3.84	2.81	2.96	2.69
dBD_2	1.68	3.25	1.99	2.31	1.98
dBE	2.00	3.41	2.46	2.62	2.52
PLdB	1.83	3.48	2.24	2.52	2.28
PNdB	1.83	3.58	2.04	2.48	2.12
PNdBM	1.62	3.37	2.00	2.33	2.03
$PNdBM_{t1}$	2.68	3.11	2.73	2.84	2.43
$PNdBM_{t2}$	2.17	3.84	2.52	2.84	2.33
$PNdB_{t2}$	2.32	3.54	2.56	2.81	2.28

Table 4 - Part 2 PAIRED-COMPARISON TEST - EPNL

Standard deviation of differences between the physical level of a standard noise and of comparison noises that are found to be subjectively equal in acceptability according to Paired-Comparison tests.

Standard 1				
	Group			
EPNL	I	II	III	Average
EdBC	6.04	7.25	6.33	6.53
EdBA	2.49	4.06	3.12	3.22
$EdBD_2$	1.70	2.95	2.49	2.37
EdBE	3.02	4.25	3.14	3.47
EPLdB	2.37	3.77	2.69	2.94
EPNdB	2.28	3.60	2.50	2.79
EPNdBM	2.09	3.45	2.20	2.58
$EPNdBM_{t1}$	1.65	2.80	1.76	2.07
$EPNdBM_{t2}$	1.89	3.36	2.04	2.43
$^*EPNdB_{t2}$	2.13	3.52	2.34	2.66

Standard 2					
	Group				
EPNL	I	II	III	Average	Average Of S1 And S2
EdBC	5.86	7.41	6.71	6.66	6.60
EdBA	2.71	3.69	3.35	3.06	3.14
$EdBD_2$	1.79	3.32	2.93	2.53	2.51
EdBE	2.95	3.85	3.59	3.46	3.47
EPLdB	2.56	3.65	3.41	3.21	3.07
EPNdB	2.23	3.49	2.79	2.84	2.81
EPNdBM	1.92	3.35	2.41	2.56	2.57
$EPNdBM_{t1}$	2.05	2.83	2.02	2.30	2.19
$EPNdBM_{t2}$	2.57	3.96	2.88	3.14	2.78
$^*DPNdB_{t2}$	2.99	2.80	3.03	2.94	2.80

*Unit used by FAA-ISO and SAE in Air Craft Noise Evaluation Procedures

Predictive Accuracy of the Physical Units. The variable of spectrum content of different noises has received the greatest research and engineering attention for purposes of noise control. Table 5 is a summary table of how well, in standard deviation terms, the various physical units of measurement predicted the results of the subjective judgment tests. It might be noted that the range of these standard errors is from about 2 dB from the best to over 6 dB for the worst; these values are similar in magnitude to those found in other well controlled comparative judgment tests that have been conducted in the past (Ref. 4).

The results given in Table 5 are in substantial agreement with previous experiments, in that the D_2 overall frequency weighting, and the PNdBM third-octave band means of frequency weighting, in general, give better predictions of the subjective judgments than do the other units of physical measurement. Also, the tone corrections, t_1 and t_2, show some utility in this regard; however, as usual with rather complex experiments of the sort involved, the effects of the tone corrections are rather small.

Of special interest is, perhaps, the finding that the units of physical measurement PLdB and dBE recently proposed by Stevens (Ref. 5) do not predict the subjective value of the noises involved in these tests any better than does dBA on the average, having average standard deviations of 2.54 for PLdB, 2.84 for dBE, and 2.78 for dBA.

In interpreting these data, it should be borne in mind that from a strictly statistical point of view a difference of about 0.25 to 0.50 dB between two standard deviations of values of the order of 2.0 dB is significant with the number of data points (a total of 126) in these tests. According to the F test of statistical significance, an F* of about 1.27 with an N,

---

*$F = \dfrac{S.D.^2 \text{ Larger}}{S.D.^2 \text{ Smaller}}$. Ref. 11

Table 5

RANK ORDERING OF STANDARD DEVIATIONS (S.D.) AVERAGED FOR MAGNITUDE-ESTIMATION AND PAIRED-COMPARISON RESULTS AND FOR THREE GROUPS OF SUBJECTS.

Max PNL			EPNL			Aver. PNL and EPNL		
Rank	Unit	S.D.	Rank	Unit	S.D.	Rank	Unit	Aver. S.D.
1.	dBD_2	1.97	1.	$PNdBM_{t1}$	1.99	1.	dBD_2	2.17
2.	PNdBM	1.97	2.	$PNdM_{t2}$	2.29	2.	PNdBM	2.20
3.	PNdB	1.99	3.	dBD_2	2.36	3.	$PNdB_{t2}$	2.20
4.	$PNdB_{t2}$	2.03	4.	$PNdB_{t2}$	2.36	4.	$PNdBM_{t2}$	2.21
5.	$PNdBM_{t2}$	2.12	5.	PNdBM	2.42	5.	$PNdBM_{t1}$	2.24
6.	PLdB	2.20	6.	PNdB	2.64	6.	PNdB	2.32
7.	dBA	2.45	7.	PLdB	2.88	7.	PLdB	2.54
8.	dBE	2.46	8.	dBA	3.11	8.	dBA	2.78
9.	$PNdBM_{t1}$	2.48	9.	dBE	3.22	9.	dBE	2.84
10.	dBC	5.79	10.	dBC	6.19	10.	dBC	5.99

number of data points, of 126 would be significant at the 95% level of confidence (Ref. 11). An F of 1.27 would be reached with a larger S.D. of 2.25, and a smaller S.D. of 2.00; a difference of 0.5 dB in the standard deviations, would be significant at a confidence level of 95% with an N of but about 42. In the present study, combining the judgments of the 14 noises made by the 3 groups of subjects under three separate test conditions (the method of paired comparison with two standards and the method of magnitude estimation) gives an N of 126. N for comparison of one Group of subjects and one test condition is, of course, 14; for one Group and two test conditions, N is 28, etc.

From a more practical point of view, it could be argued (without success with persons concerned primarily with overall environmental noise evaluation as compared to noise control at the source or in machinery design) that a difference of about 0.50 dB in these standard errors should be considered significant. The argument is based on the fact that according to normal probability statistics it is reasonable to expect that populations of noises of the types studied will be subjectively about the same when their measured levels are within plus or minus three ($\pm$ 3) or six standard deviations. Thus, the range of levels for noises of similar subjective value would be of the order of, for Max PNL, 11.82 dB (S.D. of 1.97 x 6) for the D_2 weighting and 14.70 dB (S.D. of 2.45 x 6) for the A weighting. The increased error range of 3 dBA over the dBD_2 range of expected error could be important to achieving valid design and noise control goals in some cases, inasmuch as a difference in 3 dB is a matter of 100% in sound power. This is equivalent, for example, to a doubling (or a halving) of the number of engines on an aircraft for equal subjective effect.

A more detailed statistical analysis of the Max PNL and EPNL values for the D_2 and A weighted sound levels of the noises is presented in Table 6. It is seen in Table 6 that, overall conditions: Max dBD_2 is significantly better at the 90% level of confidence than Max dBA; $EdBD_2$ is better than EdBA with a

Table 6

"F" Tests of Statistical Significance of Differences Between dBD_2 and dBA Results for Max PNL and EPNL According to Methods of Paired Comparison and Magnitude Estimation for all Three Groups of Subjects Combined. Results and Standard Deviations (S.D.) of Distributions of Physical Measures of Noises Judged to be Subjectively Equal in Noisiness.

		Max dBD_2 vs. Max dBA		$EdBD_2$ vs. EdBA		Max dBD_2 vs. $EdBD_2$		Max dBA vs. EdBA	
P.C. S1	S.D.	2.19	2.81	2.76	3.44	2.19	2.76	2.81	3.44
		F 1.65**		F 1.55**		F 1.59**		F 1.50**	
P.C. S2	S.D.	2.75	3.09	3.09	3.54	2.75	3.09	3.09	3.54
		F 1.26		F 1.31		F 1.26		F 1.31	
P.C. S1 & S2	S.D.	2.76	3.17	3.14	3.68	2.76	3.14	3.17	3.68
		F 1.32**		F 1.37**		F 1.30		F 1.35**	
M.E.	S.D.	2.68	2.81	2.82	3.37	2.68	2.82	2.81	3.37
		F 1.10		F 1.43		F 1.11		F 1.44	
P.C.S1, S2&M.E.	S.D.	2.82	3.13	3.10	3.62	2.82	3.10	3.13	3.62
		F 1.23**		F 1.36*		F 1.21		F 1.34*	

* Significant at 95% level of confidence

** Significant at 90% level of confidence

more than 95% level of confidence; Max dBA is also more accurate than EdBA; and Max dBD_2 barely misses being statistically better at the 90% level of confidence than $EdBD_2$.

Comparison of Max PNL with EPNL. One of the surprising things revealed in Tables 3, 4, 5 and 6 is that the Max PNL units of physical measurement predict so well the subjective judgments of the noises that varied so much in duration -- sounds from 4 secs. to over 18 secs. duration. Indeed, Max PNL has an edge over EPNL. Such a result is usually found when the judged noises are of comparable durations (in which case the duration effect is more or less a constant), or the subjects seem to concern themselves solely with judging the noises with respect to the peak levels reached by each of a variety of noises, rather than to judge how the longer duration noises might affect them in "real life." It has been demonstrated by some laboratory tests and in some field tests conducted with actual aircraft flyovers that similar spectra noises of longer duration are judged as less acceptable than the same noises when of shorter duration.

It should perhaps be noted that except for the two standards, S1 and S2, the various noises that differed greatly in duration were also of considerably different spectra type and that the contribution of the "skirt" energy of the noises that were of longer duration than the average would not be very large. More important, however, is probably that in spite of the instructions to "judge the whole noise" the subjects placed heavier weight in their judgments upon the "peak" levels of the noises as being, under the laboratory circumstances, the most obvious aspect of the various noises that could readily and reliably be comparatively judged and subjectively quantified. It is hypothesized that the independent variations of such factors as duration, spectral shape and complexity, and rates of the growth and decay of the noise, tend to force the subjects making judgments of a conglomeration of such differing noises, to attend primarily to such common features as the general spectral

content and peak level.

Differences Between Test Methods. In an earlier study (Ref. 1) it was found that the methods of paired comparison and magnitude estimation gave comparable results, both in terms of reliability and the general conclusions regarding the predictiveness of various physical measures of the noise. The present data, by and large, substantiate these findings, as shown in Table 7 for each group of subjects. In addition, as shown more succinctly in Table 8 when the data for all three groups are combined, there are no significant differences between the results obtained when either standard was used with the method of paired comparison or between the results obtained with the method of paried comparison and the method of magnitude estimation.

Group Differences. Inasmuch as the subjects were assigned to the three groups on essentially a random basis, consistent differences between the results for the groups would presumably be attributable to the effects of differences in the instructions to the groups and/or the order in which the tests were administered. It is seen in Table 9 that the paired comparison judgments made by Group I were generally less variable, and with some statistical significance, than those made by Groups II and III, and that Group II was slightly less consistent in their paired comparison judgments than was Group III. However, the three groups performed about equally well in their magnitude estimation judgments.

These relations are perhaps better illustrated in Table 10 where it is seen that, except for Max PNL, Group II subjects were more variable in their judgments, as predicted by the physical measurements of the noise, than were the subjects in Groups I or III at either the 5% or 10% level of confidence. Thus, it might be conjectured that the paired comparison test is adversely influenced by previous experience with a magnitude estimation test, but that the reverse is not true. It is more likely, however, that the general increase in the size of the standard deviations for the Group II paired

Table 7. "F" Tests of Statistical Significance of Differences Between Results, in Max dBD_2 and $EdBD_2$, for Methods of Paired Comparison and Magnitude Estimation Tests for Each Group of Subjects Separately.

Paired Comparison Reference Standards							
Reference Standard 1 vs Reference Standard 2							
Max dBD_2				$EdBD_2$			
Group	I	II	III	Group	I	II	III
SD_{S1}	1.75	2.41	1.15	SD_{S1}	1.70	2.95	2.49
SD_{S2}	1.64	3.05	1.99	SD_{S2}	1.79	3.35	2.93
F	1.14	1.60	2.99*	F	1.11	1.29	1.38
Magnitude Estimation vs Paired Comparison							
Magnitude Estimation vs Reference Standard S1							
Group	I	II	III	Group	I	II	III
SD_{ME}	2.20	1.53	2.03	SD_{ME}	1.89	2.00	2.26
SD_{S1}	1.75	2.41	1.15	SD_{S1}	1.70	2.95	2.49
F	1.58	2.48**	3.12*	F	1.24	2.18**	1.21
Magnitude Estimation vs Reference Standard S2							
Group	I	II	III	Group	I	II	III
SD_{ME}	2.20	1.53	2.03	SD_{ME}	1.89	2.00	2.26
SD_{S2}	1.64	3.05	1.99	SD_{S2}	1.79	3.32	2.93
F	1.80	3.97*	1.04	F	1.11	2.76*	1.68
Magnitude Estimation vs Both Standards Combined							
Group	I	II	III	Group	I	II	III
SD_{ME}	2.20	1.53	2.03	SD_{ME}	1.89	2.00	2.26
$SD_{S1,2}$	2.10	2.86	2.19	$SD_{S1,2}$	2.07	3.27	2.98
F	1.10	3.49 *	1.16	F	1.20	2.67 *	1.74

*Significant at 95% level of confidence

**Significant at 90% level of confidence

Table 8

"F" Tests of Statistical Significance of Differences Between Results in Max dBD_2 and $EdBD_2$ for Methods of Paired Comparison and Magnitude Estimation Tests for all Three Groups of Subjects Combined.

Paired Comparison Reference Standards			
Reference Standard 1 vs. Reference Standard 2			
	Max dBD_2		Max dBD_2
SD	2.19	F 1.58	2.75
	$EdBD_2$		$EdBD_2$
SD	2.76	F 1.25	3.09
Magnitude Estimation vs. Paired Comparison			
M.E. vs. P.C. Reference Standard 1			
	Max dBD_2		Max dBD_2
SD	2.68	F 1.50	2.19
	$EdBD_2$		$EdBD_2$
SD	3.14	F 1.29	2.76
M.E. vs. Reference Standard 2			
	Max dBD_2		Max dBD_2
SD	2.68	F 1.05	2.75
	$EdBD_2$		$EdBD_2$
SD	3.14	F 1.03	3.09
M.E. vs. Both Standards Combined			
		Max dBD_2	
SD	2.68	F 1.06	2.76
		$EdBD_2$	
SD	2.82	F 1.24	3.14

Table 9

"F" Tests of Statistical Significance Between Results in Max dBD_2 and $EdBD_2$, of the Different Groups of Subjects for the Methods of Paired Comparison and Magnitude Estimation.

			Group I vs Group II		Group I vs Group III		Group II vs Group III	
PC, S1	Max dBD_2	SD	1.75	2.41	1.75	1.15	2.41	1.15
		F	1.91**		2.32**		4.39*	
	$EdBD_2$	SD	1.70	2.95	1.70	2.49	2.95	2.49
		F	3.01*		2.15**		1.40	
PC, S2	Max dBD_2	SD	1.64	3.05	1.64	1.19	3.05	1.99
		F	3.46*		1.47		2.35**	
	$EdBD_2$	SD	1.79	3.32	1.79	2.93	3.32	2.93
		F	3.44*		2.68**		1.28	
PC, S1 & S2	Max dBD_2	SD	2.10	2.86	2.10	2.19	2.86	2.19
		F	1.85**		1.09		1.71**	
	$EdBD_2$	SD	2.07	3.27	2.07	2.98	3.27	2.98
		F	2.50*		2.07**		1.20	
Mag. Est.	Max dBD_2	SD	2.20	1.53	2.20	2.03	1.53	2.03
		F	2.07**		1.17		1.76	
	$EdBD_2$	SD	1.89	2.00	1.89	2.26	2.00	2.26
		F	1.12		1.43		1.28	

*Significant at 95% level of confidence

**Significant at 90% level of confidence

Table 10

"F" Tests of Statistical Significance of Differences Between Results, in Max dBD_2 and $EdBD_2$, of the Different Groups of Subjects for Methods of Paired Comparison, both Reference Standards, and Magnitude Estimation Combined.

	Group I vs Group II		Group I vs Group III		Group II vs Group III	
	Max dBD_2	Max dBD_2	Max dBD_2	Max dBD_2	Max dBD_2	Max dBD_2
SD	2.47	2.62	2.47	2.12	2.62	2.12
F	1.13		1.36		1.53 **	
	$EdBD_2$	$EdBD_2$	$EdBD_2$	$EdBD_2$	$EdBD_2$	$EdBD_2$
SD	2.43	3.00	2.43	2.74	3.00	2.74
F	1.52 **		1.27		1.99 *	

Group I - Long Form Instructions; P.C. Tests First, M.E. Second.
Group II - Short Form Instructions; M.E. Tests First, P.C. Second.
Group III - Short Form Instructions; P.C. Tests First, M.E. Second.

* Difference Significant at 95% Level of Confidence.

** Difference Significant at 90% Level of Confidence.

comparison tests, in comparison with the other results, is not significant and was due to some unidentified experimental error. In any event, the results obtained with Group I, who received the longer, more detailed instructions, were not significantly different from those found with Group III, who received the shorter set of instructions.

Classes of Noise. One of the purposes of the present tests was to determine whether or not commonly used physical measurements could be used as well for the evaluation of so-called VSTOL-type noise as for fixed wing, jet aircraft noise. The detailed data with respect to the physical units of dBA and dBD_2 for each of the noises evaluated are presented in Table 11A, B and C. Casual examination of Table 11 appears to show no striking pattern between the different types of noises and the proficiency with which their subjective ratings were predicted by these three physical units of noise measurement. However, as shown in Table 12, grouping the noise by type does reveal that possibly the subjective effect of the noise from the air-conditioner (AC) is not predicted quite as well as are the effects of the aircraft noises. Because only one air-conditioner noise was tested it is, of course, not possible to consider this finding as reliable. It is to be noted in Table 12 that the VSTOL noises and the noises from the typical jet aircraft are predicted by the physical measures about equally well.

Magnitude Scale of Perceived Noisiness. A special feature of the magnitude estimation test procedure is that a ratio scale of the subjective quantity (noisiness or unacceptability) is obtained from numerical values ascribed to the noises by the subjects, i.e., a noise given the number "twenty" presumably being twice as unacceptable to the listener as a noise given the numerical rating of 10. It has typically been found in the past that when judging non-impulsive noises, at least in the mid-to-high range of intensities, a 10 dB increase in the sound pressure level of the noise would cause a doubling in its subjective loudness or noisiness.

Table 11A

PAIRED COMPARISON--STANDARD 1

Group	Maximum dBA			Maximum dBD_2			E dBA			E dBD_2		
	I	II	III	I	II	III	I	II	III	I	II	III
Noises												
S1	81.000	81.000	81.000	88.100	88.100	88.100	87.300	87.300	87.300	94.400	94.400	94.400
S2	76.000	76.500	77.500	83.300	83.800	84.800	87.600	88.100	89.100	94.600	97.000	96.300
V1	81.500	82.000	81.000	87.200	87.700	86.700	89.800	90.300	89.300	95.600	96.100	95.100
V2	80.000	82.500	81.500	86.000	88.500	87.500	88.300	90.800	89.800	94.300	96.800	95.800
V3	81.000	85.500	82.000	85.600	90.100	86.600	92.300	96.800	93.300	96.100	100.600	97.100
V4	77.000	79.000	80.000	82.400	84.400	85.400	88.600	90.600	91.600	92.800	94.800	95.800
V5	81.000	86.000	83.500	83.600	88.600	86.100	93.000	98.000	95.500	96.000	101.000	98.500
F1	80.000	84.500	80.500	87.000	91.500	87.500	87.800	92.300	88.300	93.600	98.100	94.100
F3	81.500	88.000	84.000	85.200	91.700	87.700	91.800	98.300	94.300	96.700	103.200	89.200
F4	80.000	82.000	79.500	86.000	88.000	85.500	88.400	90.400	87.900	94.500	96.500	94.000
F5	77.000	80.500	80.500	82.400	85.900	85.900	86.300	89.800	89.800	90.900	94.400	94.400
F6	80.000	84.500	80.000	84.800	88.800	84.800	90.100	94.100	90.100	94.700	98.700	94.700
F7	76.000	77.500	76.000	84.200	85.700	84.200	84.200	85.700	84.200	92.100	93.600	92.100
AO	80.000	85.000	82.500	84.800	89.800	87.300	91.500	96.500	94.000	96.500	101.500	99.000
Mean	79.43	82.46	80.68	85.04	88.04	86.22	89.01	92.07	90.32	94.51	97.70	95.04
Standard deviation	2.00	3.35	2.14	1.75	2.41	1.15	2.49	4.06	3.12	1.70	2.95	2.49

Table 11B

PAIRED COMPARISON--STANDARD 2

	Maximum dBA			Maximum dBD_2			E dBA			E dBD_2		
Group	I	II	III	I	II	III	I	II	III	I	II	III
Noises												
S1	77.000	78.500	77.500	84.100	85.600	80.600	84.300	85.800	84.800	93.800	87.900	91.900
S2	73.500	73.500	73.500	80.800	80.800	80.800	84.600	84.600	84.600	91.800	91.800	91.800
V1	76.000	85.000	80.000	81.700	90.700	85.700	84.300	93.300	88.300	90.100	99.100	94.100
V2	76.000	80.000	76.000	82.000	86.000	82.000	84.300	88.300	84.300	90.300	94.300	90.300
V3	80.000	81.000	81.500	84.600	85.600	86.100	91.300	92.300	92.800	95.100	96.100	96.600
V4	74.000	79.500	77.000	79.400	84.900	82.400	85.600	91.100	88.600	89.800	95.300	93.800
V5	79.000	80.000	80.500	81.600	82.600	83.100	91.100	92.000	92.500	94.000	95.000	95.500
F1	78.000	85.500	79.000	85.000	92.500	86.000	85.800	93.300	86.800	91.600	99.100	92.600
F3	80.000	85.000	80.000	83.700	88.700	83.700	90.300	95.300	90.300	95.200	100.200	85.200
F4	78.000	78.000	78.500	84.000	84.000	84.500	86.400	86.400	86.300	92.500	92.500	93.000
F5	78.000	81.000	78.000	83.400	86.400	83.400	87.300	90.300	87.300	91.900	94.900	91.900
F6	76.500	82.000	80.000	81.300	86.800	84.800	86.600	92.100	90.100	91.200	96.700	94.700
F7	74.500	76.000	72.500	82.700	84.200	80.700	82.700	84.200	80.700	90.600	92.100	88.600
AO	76.500	82.500	77.500	81.300	87.300	82.300	88.000	94.000	89.000	93.000	99.000	94.000
Mean	76.93	80.54	77.96	82.54	86.15	83.26	86.6	90.21	87.60	92.23	95.51	92.43
Standard deviation	2.06	3.43	2.60	1.64	3.05	1.99	2.71	3.69	3.35	1.79	3.32	2.93

Table 11C

MAGNITUDE ESTIMATION

	Maximum dBA			Maximum dBD_2			E dBA			E dBD_2		
Group	I	II	III	I	II	III	I	II	III	I	II	III
Noises												
S1	80.000	80.000	80.000	87.100	87.100	87.100	87.300	87.300	87.300	91.400	94.400	94.400
S2	73.000	79.000	79.000	80.300	86.300	86.300	84.600	90.600	90.600	91.800	97.800	97.800
V1	77.000	81.000	77.000	82.700	86.500	82.700	85.300	89.300	85.300	91.100	95.100	91.100
V2	75.000	80.000	80.000	81.000	86.000	86.000	83.300	88.300	88.300	89.300	94.300	94.300
V3	76.000	81.000	78.000	80.600	85.600	82.600	87.300	92.300	89.300	91.100	96.100	93.100
V4	75.000	79.000	80.000	80.400	84.400	85.400	86.600	90.600	91.600	90.800	94.800	95.800
V5	76.000	82.000	81.000	78.600	84.600	83.600	88.000	94.200	93.000	91.000	97.000	96.000
F1	76.000	80.000	82.000	83.000	87.000	89.000	83.800	87.800	89.800	89.600	93.600	95.600
F3	77.000	82.000	81.000	80.700	85.700	84.700	87.300	92.300	91.300	92.200	97.200	96.200
F4	75.000	81.000	78.000	81.000	87.000	84.000	83.400	89.400	86.400	89.500	95.500	92.500
F5	74.000	79.000	80.000	79.400	84.400	85.400	83.300	88.300	89.300	87.900	92.900	93.900
F6	74.000	78.000	79.000	78.800	82.800	83.800	84.100	88.100	89.100	88.700	92.700	93.700
F7	71.000	74.000	73.000	79.200	82.200	81.200	79.200	82.200	81.200	87.200	90.600	89.100
AO	78.000	80.000	79.000	82.200	84.800	83.800	89.500	91.500	90.500	94.500	96.500	95.500
Mean	75.50	79.71	79.07	81.07	85.31	84.69	85.4	89.44	88.79	90.4	94.89	94.21
Standard deviation	2.21	2.02	2.20	2.20	1.53	2.03	2.65	2.91	3.01	1.89	2.00	2.26

Table 12

Average Levels in dB for the units dBA and dBD_2 for 4 classes of Noises Judged to be Equally Acceptable in the Paired-Comparison and Magnitude-Estimation Tests.

	MAX PNL		EPNL	
	Max dBA	Max dBD_2	EdBA	$EdBD_2$
Av. VSTOL	79.70	84.56	90.0	94.58
Av. Fixed Wing Jet	79.01	84.85	88.19	93.30
Air-Conditioner	80.11	84.84	91.61	96.61
Pink Noise	77.64	84.16	86.84	93.93
Mean of all 14 Noises	79.14	84.70	88.83	94.10

Difference From Means of All 14 Noises

	MAX PNL		EPNL	
	Max dBA	Max dBD_2	EdBA	$EdBD_2$
Av. VSTOL	+0.56	-0.14	+1.17	+0.48
Av. Fixed Wing Jet	-0.13	+0.15	-0.64	-0.8
Air Conditioner	+0.97	+0.14	+2.78	+2.51
Pink Noise	-1.5	-0.54	-1.99	-0.17

The results of magnitude estimation tests are tabulated for the highest and lowest levels of each noise in Table 13. Fig. 6 is a summary plot of the averages for all the noises when presented at three different levels of intensity. It is seen in Fig. 6 that the traditional doubling in the perceived noisiness occurs with a 10 dB increase in intensity.

Table 13

MAGNITUDE ESTIMATES, AND RATIOS, FOUND FOR NOISES WHEN AT DIFFERENT LEVELS OF INTENSITY

Group I							Group II							Group III						
Noise	Lowest dBA Level	Highest dBA Level	dB Diff. (H-L)	M.E. Lowest Level	M.E. Highest Level	M.E. Ratio (H-L)	Noise	Lowest dBA Level	Highest dBA Level	dB Diff. (H-L)	M.E. Lowest Level	M.E. Highest Level	M.E. Ratio (H-L)	Noise	Lowest dBA Level	Highest dBA Level	dB Diff. (H-L)	M.E. Lowest Level	M.E. Highest Level	M.E. Ratio (H-L)
S1	74	84	10	7.7	18.5	2.4	S1	74	84	10	10.9	19.5	1.79	S1	74	84	10	9.3	19.7	2.12
S2	66	82	16	6.5	25.0	3.85	S2	66	82	16	6.5	20.7	3.18	S2	66	82	16	6.2	21.2	3.42
V1	72	88	16	6.5	22.3	3.43	V1	72	88	16	7.0	17.6	2.51	V1	72	88	16	10.3	21.2	2.06
V2	70	86	16	6.6	30.3	4.59	V2	70	86	16	6.8	23.3	3.43	V2	70	86	16	7.3	29.8	4.08
V3	72	88	16	7.2	34.7	4.82	V3	72	88	16	6.5	20.7	3.18	V3	72	88	16	8.7	28.3	3.25
V4	70	86	16	9.9	34.2	3.45	V4	70	86	16	7.4	22.7	3.07	V4	70	86	16	6.9	26.4	3.83
V5	70	86	16	9.0	27.7	3.08	V5	70	86	16	8.7	17.8	2.05	V5	72	86	16	7.9	20.0	2.53
F1	72	88	16	6.8	31.7	4.59	F1	72	88	16	6.5	24.6	3.78	F1	72	88	16	8.0	27.3	3.41
F3	72	88	16	7.4	25.0	3.38	F3	72	88	16	8.3	18.5	2.23	F3	72	88	16	6.9	22.7	3.29
F4	72	88	16	10.2	33.3	3.26	F4	72	88	16	8.1	23.8	2.94	F4	72	88	16	8.8	30.8	3.50
F5	70	86	16	9.1	29.7	3.26	F5	70	86	16	8.0	21.8	2.72	F5	70	86	16	10.1	22.7	2.25
F6	72	88	16	8.3	34.7	4.18	F6	72	88	16	9.2	26.5	2.88	F6	72	88	16	8.7	28.5	3.27
F7	66	82	16	5.6	28.8	5.14	F7	66	82	16	7.2	25.6	3.55	F7	66	82	16	5.9	23.0	3.89
AC	70	86	16	8.2	28.7	3.50	AC	70	86	16	8.0	21.3	2.66	AC	70	86	16	8.4	26.5	3.15

Ratios Between ME's Given to Different
Levels of Same Noise

	dB Diff. Between Two Levels of Same Noise		
	8 dB	10 dB	16 dB
Group I	1.85	2.4	4.1
Group II	1.78	1.79	3.01
Group III	1.80	2.12	3.27
Average	1.81	2.10	3.46

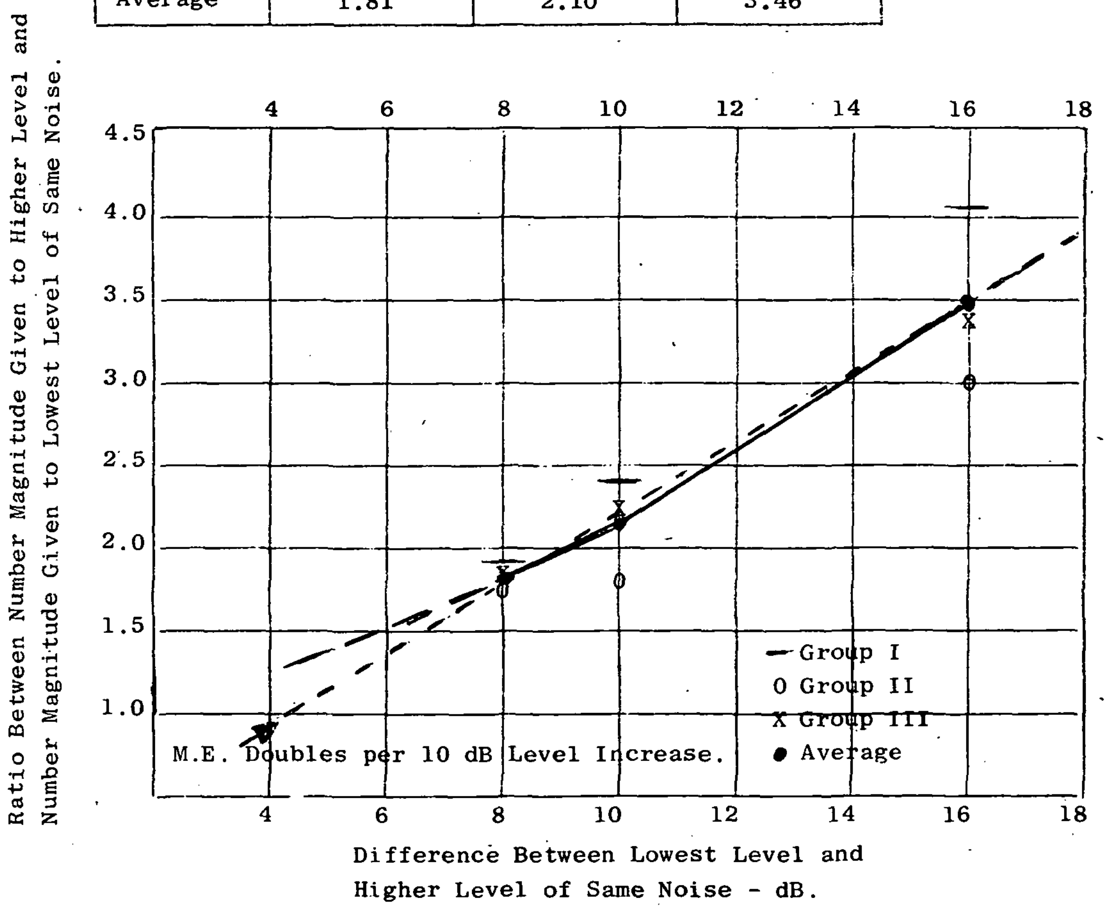


FIGURE 6. Showing ratio of magnitude estimations of subjective noisiness or unacceptableness of a noise as function of difference in physical level of two presentations of same noise.

CONCLUSIONS

On the basis of the present experimental results and related considerations presented in the discussion, it is concluded that:

1. The standard deviations with which the various units of physical noise measurement predicted the subjective judgments of perceived noisiness or unwantedness were approximately the same for both the methods of magnitude estimation and paired comparison.

2. The relative accuracy with which the physical units of measurement predicted the subjective judgments was as good for the maximum level reached by each noise as for so-called effective or time integrated levels of the noise measurements. This perhaps somewhat anomalous finding is ascribed to possible difficulties subjects have in such laboratory tests in attending simultaneously to more than one major and variable physical aspect of noises when several of these aspects are non-systematically varied among the noise stimuli.

3. Statistically significant and often practically significant differences were found in the proficiency with which the different frequency weighting procedures, both overall and one-third octave band, predicted the subjective judgment test results. Frequency weightings D_2 and PNdBM were consistently better by a small but probably a practically significant amount (about 3.5 dB over a range of $\pm$ 3 standard deviations) than the overall frequency weightings of A and E and the one-third octave band procedure of PLdB.

4. The physical units of dBA, dBD_2 and $PNdBM_{tl}$ predicted the subjective judgments of the VSTOL type aircraft noises as well as they predicted the judgments of the noises from the typical fixed wing jet aircraft.

5. The scale of perceived noisiness as determined from the magnitude estimation tests was consistent with that typically found in the past for both loudness and perceived noisiness, namely, a doubling of perceived magnitude for each increase of 10 dB in intensity.

REFERENCES

1. T. R. Clarke, and K. D. Kryter, "The Methods of Paired Comparisons and Magnitude Estimation in Judging the Noisiness of Aircraft," NASA Report No. CR-2107 (August 1972).

2. The Boeing Company, "Study and Development of Turbofan Nacelle Modifications to Minimize Fan-Compressor Noise Radiation," Vol. VII. Subjective Evaluation Tests. NASA Report No. CR-1717, 1971, prepared by The Boeing Company under Contract No. NAS1-7129 for NASA Langley Research Center.

3. L. E. Langdon, R. F. Gabriel, and A. H. Marsh, "Investigation of DC-8 Nacelle Modifications to Reduce Fan-Compressor Noise in Airport Communities," Part VI - Psychoacoustic Evaluation. NASA Report No. CR-1710, 1970, prepared by the McDonnell Douglas Corporation under Contract No. NAS1-7130 for NASA Langley Research Center.

4. K. D. Kryter, The Effects of Noise on Man, Academic Press, New York, New York, (1970).

5. S. S. Stevens, Perceived Level of Noise by Mark VIII and Decibels (E) J. Acoust. Soc. Am. 51, 573-601, (1972).

6. American Standard Specification for General Purpose Sound Level Meters. ASAS1.4-1961, American National Standards Institute, New York.

7. K. D. Kryter, P. J. Johnson, and J. R. Young, "Judgment Tests of Flyover Noise from Various Aircraft," NASA Report No. CR-1635, Stanford Research Institute, National Aeronautics and Space Administration, Washington, D.C. (1969).

8. K. D. Kryter, "A Note of the Quantity (Effective) Perceived Noisiness and Units of Perceived Noise Level," J. Sound and Vibration, Vol. 25, No. 3, 383-393.

9. D. E. Broadbent, and D. W. Robinson, "Subjective Measurements of the Relative Annoyance of Simulated Sonic Bangs and Aircraft Noise," J. Sound and Vibration, Vol. 1, No. 2, 162-174 (1964).

10. J. E. Parnell, D. C. Nagel, and H. J. Parry, "Growth of Noisiness for Tones and Bands of Noise at Different Frequencies," Technical Report FAA-DS-67-21, 1967, prepared under Contract No. FA65WA-1180 by Bolt, Beranek and Newman, Inc., for the Federal Aviation Administration, Washington, D.C.

11. A. M. Mood, "Introduction to the Theory of Statistics," McGraw-Hill, New York (1950).

BRIEF INSTRUCTIONS - PAIRED COMPARISON

You will hear a series of sounds from aircraft. The sounds will occur in "pairs" and your task is to judge which sound in each pair you think would be more unacceptable, bothersome or annoying to you.

After you have heard each pair of sounds, please quickly decide which of the two you feel would be more unacceptable or annoying to you. If you think the first sound of a pair to be more unacceptable, circle A for that particular pair. If you think the second sound in the pair would be more unacceptable to you than the first, circle B. If you feel that there is absolutely no real difference in terms of acceptability of the two sounds, please circle either A or B, giving the best guess you can.

An announcement of the item number will be made before each <u>5th pair of sounds is to occur</u>. The sounds of a pair will be separated by a few seconds.

LONG INSTRUCTIONS - PAIRED COMPARISON

The primary purpose of the tests being conducted is to determine, if possible, how people feel about the relative unacceptability or annoyingness of one type or level of aircraft noise when compared with a second type or level of aircraft noise.

You will hear a series of sounds from aircraft. The sounds will occur in "pairs" and your task is to judge which sound in each pair you think would be more unacceptable, bothersome or annoying to you if heard in or near your home during the day and/or evening when you are engaged in typical, awake activities. Judge how the whole, entire sound or noise would affect you - not just the peak level but the whole noise from beginning to end as though you were in your home.

After you have heard each pair of sounds, please quickly decide which of the two you feel would be more unacceptable or annoying to you. If you think the first sound of a pair to be more unacceptable, circle A for that particular pair. If you think the second sound in the pair would be more unacceptable to you than the first, circle B.

Please concentrate on the judgment at hand and give an answer even though the two sounds may seem approximately equal in acceptability or unacceptability to you. If you feel that there is absolutely no real difference in terms of acceptability of the two sounds, please circle either A or B, giving the best guess you can, and put a question mark after that pair.

There are no "right" or "wrong" answers, nor do we expect people to agree with each other. We are interested in how you feel about the sounds and how people differ in their judgments of these aircraft sounds in their entirety in or near your home.

An announcement of the item number will be made before each <u>5th pair of sounds is to occur</u>. The sounds of a pair will be separated by a few seconds. During the test period, which will be approximately 20 minutes, please remain quiet and attentive. Give us your best judgment and imagine, if you will, that you are listening to these sounds in or near your own home.

LONG INSTRUCTIONS - MAGNITUDE ESTIMATION

The primary purpose of the tests being conducted is to determine, if possible, how people feel about the relative unacceptability or annoyingness of different types of aircraft noise.

You will hear a series of sounds from aircraft. Your task is to judge how unacceptable, bothersome or annoying each sound would be to you if heard in or near your home during the day and/or evening when you are engaged in typical, awake activities. Judge how the whole, entire sound or noise would affect you - not just the peak level but the whole noise from beginning to end as though you were in your own home.

First, we will produce a sound whose noisiness score is 10. This will be the first sound after the announcement "begin test." Use that sound as a standard, and judge each succeeding sound in relation to that standard. For example, if a sound seems <u>twice</u> as noisy or annoying or unacceptable as the standard, you will write 20 in the appropriate box on the answer sheet. If it seems only <u>one-quarter</u> as noisy, write 2.5. If it seems three times as noisy, write 30; <u>one-half</u> as noisy, write 5, and so on.

Please concentrate on the judgment at hand and give an answer that tells how strong the annoyance seems to you. There are no "right" or "wrong" answers, nor do we expect people to agree with each other. We are interested in how you feel about the sounds and how people differ in their judgments of these aircraft sounds in their entirety in or near your home.

An announcement of the item number will be made before each 5th sound. The sounds will be separated by a few seconds. During the test period, which will be approximately 20 minutes, please remain quiet and attentive. Give us your best judgment and imagine, if you will, that you are listening to these sounds in or near your own home.

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons. Also includes conference proceedings with either limited or unlimited distribution.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include final reports of major projects, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546